

WE CLAIM:

*See A1*

1. A method for extracting data from a network, comprising:
  - (a) creating a database-structured query;
  - (b) determining a web domain address on the network from which to extract the data, the web domain address having content; and
  - (c) extracting data from the determined web domain address based on the database-structured query.
2. The method of Claim 1, wherein creating the database-structured query, further comprises, including a network address within the database-structured query indicating a starting point.
3. The method of Claim 2, wherein the web domain address, includes at least one universal resource locator (URL).
4. The method of Claim 2, wherein the web domain address, further comprises, following links contained within the web domain until the links have been exhausted or following the links until a predetermined limit is reached.
5. The method of Claim 1, wherein creating the database-structured query, further comprises, creating a regular expression within the database-structured query used to determine the data to extract.
6. The method of Claim 5, wherein extracting data from the determined web domain address based on the database-structured query, further comprises, matching a plurality of patterns contained within the regular expression to the content to determine the data to extract.

*a  
cont*

7. The method of Claim 1, wherein creating the database-structured query, further comprises, creating a conditional expression within the database-structured query describing how to scan the content for the data to extract.
8. The method of Claim 1, wherein the extracting data from the determined web domain, further comprises:
  - (b) retrieving content from the web domain address;
  - (c) reducing the retrieved content to a region of interest; and
  - (d) searching the region of interest for the data matching a predetermined regular expression.
9. The method of Claim 8, wherein extracting the data from the determined web domain, further comprises, storing the data matching the predetermined regular expression.
10. The method of Claim 9, wherein extracting the data from the determined web domain, further comprises, reshaping the stored data by arranging the stored data for at least one data analysis software program.
11. A computer-readable medium having computer-executable instructions for extracting data from a network comprising:
  - (a) creating a database-structured query including a web domain address used for locating content;
  - (b) locating the content based on the web domain address; and
  - (c) extracting data based on the database-structured query from the located content.
12. The computer-readable medium of Claim 11, wherein the database-structured query, further comprises, a network address included within the database-structured query indicating a starting point.

*A  
CDN*

13. The computer-readable medium of Claim 12, wherein the network address, further comprises at least one universal resource locator (URL).
14. The computer-readable medium of Claim 11, wherein the web domain address, further comprises, links contained within the web domain to be followed until the links have been exhausted or until a predetermined limit is reached.
15. The computer-readable medium of Claim 11, wherein the database-structured query, further comprises, a regular expression within the database-structured query used to determine the data to extract.
16. The computer-readable medium of Claim 15, wherein the regular expression within the database-structured query, further comprises, a plurality of patterns used to determine the data to extract from the web domain address having content.
17. A system for extracting data from a network comprising:
  - (a) a client computer system having a client network connection to the network and communicating with a server computer system, the client creating a database-structured query;
  - (b) the server computer system having a server network connection to the network and communicating with the client computer system, the server determining a web domain address from which to extract the data from based on the database-structured query.
18. The system of Claim 17, wherein the database-structured query, further comprises, a network address within the database-structured query indicating a starting point.

*cont*

19. The system of Claim 18, wherein the database-structured query, further comprises, a regular expression within the database-structured query used to determine the data to extract.

20. The system of Claim 19, wherein the regular expression within the database-structured query, further comprises, a plurality of patterns used to determine the data to extract from the web domain address having content.

21. The system of Claim 17, further comprising an editor for creating a template of regular expressions used to extract the data.

22. The system of Claim 17, further comprising at least one data extraction engine to extract the data.

23. The system of Claim 22, wherein the data extraction engine is a web crawler.